(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101246690 B
(45)授权公告日 2011.07.13

(21)申请号 200810008294.7

(22)申请日 2008.02.15

(30)优先权数据
2007-035410 2007.02.15 JP

(73)专利权人 索尼株式会社
地址 日本东京都

(72)发明人 难波隆一 安部素嗣 井上晃
东山惠祐 高桥秀介 西口正之

(74)专利代理机构 北京东方亿思知识产权代理有限责任公司 11258
代理人 董方源

(51)Int.Cl.
G10L 21/02(2006.01)
H04R 3/00(2006.01)
H04N 5/225(2006.01)

(56)对比文件
JP 特开 2006-154314 A,2006.06.15,
JP 特开 2005-236852 A,2005.09.02,
CN 1638532 A,2005.07.13,
CN 1831554 A,2006.09.13,
WO 2006/043413 A1,2006.04.27,

审查员 王馨宁

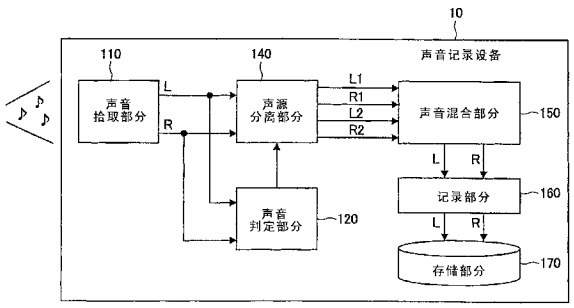
权利要求书 2 页 说明书 15 页 附图 13 页

(54)发明名称

声音处理设备及声音处理方法

(57)摘要

本发明提供了声音处理设备、声音处理方法及程序。一种声音处理设备包括：声音判定部分，用来基于特定声源的位置信息判定输入声音是否包括从该声源发出的第一声音；声音分离部分，用来在声音判定部分判定输入声音包括第一声音的情况下，将输入声音分离成第一声音和从不同于特定声源的声源发出的第二声音；以及声音混合部分，用来以预定的音量比混合由声音分离部分分离的第一声音和第二声音。



CN 101246690 B

1. 一种声音处理设备，包括：

声音判定部分，用来基于输入声音的音量和品质中的至少一个和声源的位置信息来判定输入声音是否包括从特定声源发出的第一声音；

声音分离部分，用来在所述声音判定部分判定所述输入声音包括所述第一声音的情况下，将所述输入声音分离成所述第一声音和从不同于所述特定声源的声源发出的第二声音；以及

声音混合部分，用来以预定的音量比混合由所述声音分离部分分离的所述第一声音和所述第二声音。

2. 如权利要求 1 所述的声音处理设备，其中

所述特定声源位于离所述输入声音的拾取位置规定距离之内。

3. 如权利要求 2 所述的声音处理设备，其中

所述第一声音包括由用于拾取所述输入声音的设备的操作者引起的声音，并且

所述第二声音包括从拾取目标发出的声音。

4. 如权利要求 3 所述的声音处理设备，还包括：

拍摄部分，用来拍摄视频，其中

所述声音判定部分包括位置信息计算部分，所述位置信息计算部分用来基于输入声音中包括的从一个或多个声源发出的声音的音量和相位中的至少一个来计算声源的位置信息，并且若所述位置信息计算部分计算出所述输入声音的声源位置在所述拍摄部分的拍摄方向后方且所述输入声音匹配或近似于人声，则所述声音判定部分判定所述输入声音包括所述从特定的声源发出的第一声音。

5. 如权利要求 3 所述的声音处理设备，其中

若所述输入声音的声源位置在离拾取位置规定距离之内，所述输入声音包括冲激声，且所述输入声音的音量高于过去的平均音量，则所述声音判定部分判定所述输入声音包括所述从特定的声源发出的第一声音。

6. 如权利要求 1 所述的声音处理设备，包括：

多个拾取部分，用来拾取所述输入声音；以及

记录部分，用来将所述声音混合部分混合的混合声记录到存储器中。

7. 如权利要求 4 所述的声音处理设备，包括：

存储器，用来存储所述输入声音；以及

再现部分，用来再现所述存储器中存储的输入声音并将所述输入声音输出到所述声音判定部分和所述声音分离部分中的至少一个。

8. 如权利要求 1 所述的声音处理设备，包括：

音量校正部分，用来根据对所述输入声音的音量进行了校正的情况下的校正度来反向校正所述声音分离部分分离的第二声音的音量。

9. 一种声音处理设备，包括：

声音分离部分，用于分离输入声音；

声音判定部分，用于基于输入声音的音量和品质中的至少一个和声源的位置信息来判定所述声音分离部分分离的声音是否包括从特定声源发出的第一声音；以及

声音混合部分，用于以预定混合比来混合所述声音分离部分分离的所述第一声音以及

从不同于所述特定声源的声源发出的第二声音。

10. 一种用于声音处理的方法，包括：

声音判定步骤，用来基于输入声音的音量和品质中的至少一个和声源的位置信息来判定输入声音是否包括从特定声源发出的第一声音；

声音分离步骤，用来在所述声音判定步骤判定所述输入声音包括所述第一声音的情况下，将所述输入声音分离成所述第一声音和从不同于所述特定声源的声源发出的第二声音；以及

声音混合步骤，用来以预定的音量比混合由所述声音分离步骤分离的所述第一声音和所述第二声音。

11. 如权利要求 10 所述的方法，还包括：

拍摄步骤，用来拍摄视频，其中

所述声音判定步骤包括位置信息计算步骤，所述位置信息计算步骤用来基于输入声音中包括的从一个或多个声源发出的声音的音量和相位中的至少一个来计算声源的位置信息，并且若所述位置信息计算步骤计算出所 述输入声音的声源位置在所述拍摄步骤的拍摄方向后方且所述输入声音匹配或近似于人声，则所述声音判定步骤判定所述输入声音包括所述从特定的声源发出的第一声音。

12. 如权利要求 10 所述的方法，其中

若所述输入声音的声源位置在离拾取位置规定距离之内，所述输入声音包括冲激声，且所述输入声音的音量高于过去的平均音量，则所述声音判定步骤判定所述输入声音包括所述从特定的声源发出的第一声音。

声音处理设备及声音处理方法

技术领域

[0001] 本发明涉及声音处理 (sound processing) 设备、声音处理方法及程序。

背景技术

[0002] 当今，能够记录对象的视频和对象发出的声音的视频 / 声音记录设备被广泛使用。视频 / 声音记录设备的操作者能够通过操纵视频 / 声音记录设备上的操作装置来调整视频 / 声音记录设备的拍摄方向或者放大或缩小对象的图像。

[0003] 音量随着与声源距离增大而减小。因此，在视频 / 声音记录设备中，由视频 / 声音记录设备的操作者引起的诸如操作者的语音或操作装置的操作声之类的声音可能无意中以比对象发出的声音更高的音量被记录下来。

[0004] 第 2005-341073 号日本未实审专利申请公布公开的声音处理设备用于记录抑制了由操作者引起的声音的音量的声音。具体而言，该声音处理设备包括五个定向麦克风，左前、右前、左后、右后，还有一个可拆卸。因此，位于后部中央的操作者的语音实际上不被左前、右前、左后和右后麦克风中的任一个拾取，它可根据需要或目的由可拆卸的麦克风拾取。

[0005] 第 2006-154314 号日本未实审专利申请公布公开了一种使用基于 ICA（独立成分分析）的 BSS（盲声源分离）来对包括来自多个声源的声音的混合声中的来自一个或多个声源的信号进行分离的技术。

发明内容

[0006] 但是，相关技术中的视频 / 声音记录设备需要包括大量麦克风，这导致了很大的硬件尺寸。另外，相关技术中的视频 / 声音记录设备使用麦克风的方向性来选择操作者的声音，这对操作者的位置进行了约束。

[0007] 鉴于前述问题，需要一种能够通过调整从特定的声源发出的声音占全部声音的音量比例来记录声音的新的改进的声音处理设备、声音处理方法和程序。

[0008] 根据本发明的实施例，提供了一种声音处理设备，其包括 ：声音判定部分，用来基于输入声音的音量和品质中的至少一个和声源的位置信息来判定输入声音是否包括从特定的声源发出的第一声音 ；声音分离部分，用来在声音判定部分判定输入声音包括第一声音的情况下，将输入声音分离成第一声音和从不同于特定声源的声源发出的第二声音 ；以及声音混合部分，用来以预定的音量比例混合由声音分离部分分离的第一声音和第二声音。

[0009] 在该配置中，声音分离部分分离输入声音中包括的从特定的声源发出的第一声音，声音混合部分以第一声音的音量比例低于输入声音中第一声音的音量比例的方式来混合第一声音以及作为输入声音中包括的另一个声音的第二声音。因此，若输入声音中从特定的声源发出的第一声音的音量过高，则声音混合部分可以生成第二声音的音量比例高于输入声音中第二声音的音量比例的混合声。因此，声音处理设备可以防止第二声音不当地

掩埋在第一声音中。

[0010] 另外，声音混合部分可按如下方式来混合从附近发出的第一声音以及作为输入声音中包括的另一个声音的第二声音，即以第一声音的音量比例高于输入声音中第一声音的音量比例的方式。在该配置中，可以在希望拾取来自声音拾取者的声音时增强从声音拾取者发出的第一声音。若声音判定部分判定输入声音不包括第一声音，则声音分离部分可不分离输入声音。

[0011] 特定的声源可位于离输入声音的拾取位置规定距离之内。换言之，第一声音可从离输入声音的拾取位置规定距离内的位置发出。因为音量随离开声源的距离增大而减小，所以从拾取位置附近的声源发出的声音很可能以更高的音量被拾取为输入声音。因此，声音混合部分可以抑制输入声音的拾取位置附近的第一声音的音量比例，并校正由拾取位置和声源之间距离的不同造成的音量不平衡关系。

[0012] 第一声音可包括由拾取输入声音时使用的设备的操作者引起的声音，第二声音可包括从拾取目标发出的声音。在该配置中，可以抑制从操纵位于拾取输入声音时使用的设备附近的设备的操作者发出的第一声音的音量比例，并防止从拾取目标发出的第二声音不当地掩埋在第一声音中。

[0013] 声音判定部分可基于输入声音的音量和品质中的至少一个来判定输入声音是否包括第一声音。声音判定部分可基于输入声音的音量或相位来估计输入声音的声源位置信息或者输入声音中包括的从一个或多个声源发出的每个声音的声源位置信息。

[0014] 声音处理设备还可包括用来拍摄视频的拍摄部分，且声音判定部分可包括位置信息计算部分，该部分用来基于输入声音中包括的从一个或多个声源发出的声音的音量和相位中的至少一个来计算声源的位置信息，并且若位置信息计算部分计算出输入声音的声源位置在拍摄部分的拍摄方向后方且输入声音匹配或近似于人声，则声音判定部分判定输入声音包括从特定的声源发出的第一声音。操作者经常从拍摄部分的拍摄方向后方操纵声音处理设备。因此，若输入声音的声源位置在拍摄部分的拍摄方向后方且输入声音匹配或近似于人声，则声音判定部分可以判定输入声音主要包括操作者的声音作为第一声音。因此可以获得操作者声音的音量比例被声音混合部分减小的混合声。

[0015] 若输入声音的声源位置在离拾取位置规定距离之内，输入声音包括冲激声，且输入声音的音量高于过去的平均音量，则声音判定部分可以判定输入声音包括从特定的声源发出的第一声音。当输入声音拾取设备的操作者操纵设备按钮或改变设备的手持方式时，很可能出现啪嗒或砰之类的冲激声。另外，由于这类冲激声在设备中生成，因此很可能以相对较高的音量被拾取。因此，若输入声音的声源位置在离声音拾取位置规定距离之内，输入声音包括冲激声，且输入声音的音量高于过去的平均音量，则声音判定部分可以判定：输入声音主要包括由操作者的动作引起的噪声作为第一声音。因此可以获得由操作者的动作造成的噪声的音量比例被声音混合部分减小的混合声。

[0016] 声音处理设备可包括用来拾取输入声音的多个拾取部分，和用来将声音混合部分混合的混合声记录到存储器中的记录部分。在该配置中，记录部分将第一声音的音量比例低于输入声音中第一声音的音量比例的混合声记录到存储器中。这允许重放设备重放第一声音的音量比例被适当调整的混合声，而不需在重放设备上安装专门的音量校正功能。

[0017] 声音处理设备可包括：存储器，用来存储输入声音；以及再现部分，用来再现存储

器中存储的输入声音并将输入声音输出到声音判定部分和声音分离部分中的至少一个。在该配置中，声音判定部分和声音分离部分基于从再现部分输入的输入声音来生成混合声并将混合声作为再现声音输出。这实现了第一声音的音量比例被适当调整的混合声的重放，而不需在将输入声音记录到存储器的记录设备上安装专门的音量校正功能。

[0018] 声音处理设备可包括音量校正部分，用来根据对输入声音的音量进行了校正的情况下的校正度来反向校正声音分离部分分离的第二声音的音量。例如，若输入音量因为第一声音的音量过大而被整体抑制，则第二声音的音量也被相应地抑制。音量校正部分可以根据抑制输入音量的程度来增大第二声音的音量，从而避免第二声音过小。

[0019] 根据本发明的另一实施例，提供了一种声音处理设备，其包括：声音分离部分，用于分离输入声音；声音判定部分，用于基于输入声音的音量和品质中的至少一个和声源的位置信息来判定声音分离部分分离的声音是否包括从特定的声源发出的第一声音；以及声音混合部分，用于以预定混合比来混合声音分离部分分离的第一声音以及从不同于特定声源的声源发出的第二声音。

[0020] 根据本发明的另一实施例，提供了使得计算机充当如下声音处理设备的程序，所述声音处理设备包括：声音判定部分，用来基于声源的位置信息来判定输入声音是否包括从特定的声源发出的第一声音；声音分离部分，用来在声音判定部分判定输入声音包括第一声音的情况下，将输入声音分离成第一声音和从不同于特定声源的声源发出的第二声音；以及声音混合部分，用来以预定的音量比混合由声音分离部分分离的第一声音和第二声音。

[0021] 上述程序可使得包括 CPU、ROM、RAM 等的计算机的硬件资源执行位置信息计算部分、声音判定部分和声音分离部分的功能。因此可以使得实现该程序的计算机充当上述声音处理设备。

[0022] 声音判定部分可基于声源的位置信息、输入声音的音量和品质中的至少一个来判定输入声音是否包括第一声音。

[0023] 该程序还可包括用来拍摄视频的拍摄部分，且声音判定部分可包括位置信息计算部分，所述部分用来基于输入声音中包括的从一个或多个声源发出的声音的音量和相位中的至少一个来计算声源的位置信息，并且若位置信息计算部分计算出输入声音的声源位置在拍摄部分的拍摄方向后方且输入声音匹配或近似于人声，则声音判定部分判定输入声音包括从特定的声源发出的第一声音。

[0024] 若输入声音的声源位置在离拾取位置规定距离之内，输入声音包括冲激声，且输入声音的音量高于过去的平均音量，则声音判定部分可以判定输入声音包括从特定的声源发出的第一声音。

[0025] 根据本发明的另一实施例，提供了一种声音处理方法，包括以下步骤：基于输入声音的音量和品质中的至少一个和声源位置信息来判定输入声音是否包括从特定的声源发出的第一声音；若判定输入声音包括第一声音，则将输入声音分离成第一声音和从不同于特定声源的声源发出的第二声音；以及以预定的音量比混合彼此分离开的第一声音和第二声音。

[0026] 根据本发明的上述实施例，可以在适当调整了从特定声源发出的声音占全部声音的音量比例后输出或记录声音。

附图说明

[0027] 图 1 是示出使用根据本发明第一实施例的声音记录设备的场景示例的图示；

[0028] 图 2A 至 2C 是示出由普通声音记录方法记录的声音的时间域幅度的图示；

[0029] 图 3 是示出作为根据本发明第一实施例的声音处理设备示例的声音记录设备的配置的功能框图；

[0030] 图 4 是示出声音判定部分的配置的功能框图；

[0031] 图 5 是示出基于两个输入声音之间的相位差来估计输入声音的声源位置的方法的图示；

[0032] 图 6 是示出基于三个输入声音之间的相位差来估计输入声音的声源位置的方法的图示；

[0033] 图 7 是示出基于两个输入声音的音量来估计输入声音的声源位置的方法的图示；

[0034] 图 8 是示出基于三个输入声音的音量来估计输入声音的声源位置的方法的图示；

[0035] 图 9 是示出声音记录设备和操作者之间的位置关系的图示；

[0036] 图 10 是示出根据本发明的第一实施例在声音记录设备中执行的声音处理方法的流程的流程图；

[0037] 图 11 是示出根据本发明第二实施例的声音重放设备的配置的功能框图；

[0038] 图 12 是示出根据本发明第三实施例的声音重放设备的配置的功能框图；

[0039] 图 13 是相互对比地示出 AGC 实施之前声音的音量和 AGC 实施之后声音的音量的图示。

具体实施方式

[0040] 下面，将参考附图来详细描述本发明的优选实施例。注意，在本说明书及附图中，实际上具有相同的功能和结构的结构元素以相同的参考符号来表示，且这些结构元素的重复性描述被省略。

[0041] 【第一实施例】

[0042] 下面描述根据本发明第一实施例的声音记录设备 10。在该实施例中，参考图 1 和图 2A 至 2C 来描述使用声音记录设备 10 的示例性场景，然后参考图 3 至 10 来描述声音记录设备 10 的配置及操作。

[0043] 图 1 是示出使用本实施例的声音记录设备 10 的场景示例的图示。在图 1 的示例中，作为对象的孩子站在学校前面，手持具有视频拍摄功能的声音记录设备 10 的操作者将声音记录设备 10 瞄准该对象。

[0044] 响应于操作者的呼唤“喂”，对象应答“嗨”。具有视频拍摄功能的声音记录设备 10 将来自操作者的呼唤“喂”和来自对象的应答“嗨”连同对象的视频一起记录。下文中参考图 2A 至 2C 来描述由普通声音记录方法记录的声音。

[0045] 图 2A 至 2C 是示出由普通声音记录方法记录的声音的时间域幅度的图示。若声源被假定为点声源，则拾取的音量与声源和声音拾取位置之间的距离的平方成反比。于是，拾取的音量随着拾取位置远离声源而减小。因此，来自靠近拾取位置的操作者的呼唤“喂”被拾取为具有图 2A 所示幅度的声音。

[0046] 另一方面，来自离拾取位置比操作者远的对象的应答“嗨”被拾取为具有的幅度如图 2B 所示比操作者的语音小的声音。这种情况下，普通声音记录方法记录如图 2C 所示的声音，其中仅仅简单地将操作者的呼唤“喂”和对象的应答“嗨”相互叠加。

[0047] 但是，在图 2C 所示的声音中，操作者的呼唤“喂”过于显著以致对象的应答“嗨”被不当地掩埋其中。另外，操作者造成的操作噪声以相对高于来自对象的声音的音量级 (level) 被记录。于是，从对象发出的声音被操作者引起的声音掩盖，因此经常无法以操作者想要的适当音量平衡记录来自对象的声音。

[0048] 鉴于上述问题，发明了根据本实施例的声音记录设备 10。本实施例的声音记录设备 10 抑制由操作者引起的声音的音量比例并以适当的音量平衡记录来自对象的声音和由操作者引起的声音。下面详细描述声音记录设备 10 的配置和操作。

[0049] 图 3 是示出作为根据本实施例的声音处理设备示例的声音记录设备 10 的配置的功能框图。声音记录设备 10 包括声音拾取部分 110、声音判定部分 120、声源分离部分 140、声音混合部分 150、记录部分 160 和存储部分 170。虽然图 1 说明了作为声音记录设备 10 的示例的视频摄像机，但是声音记录设备 10 不限于视频摄像机，它可以是诸如 PC（个人电脑）、移动电话、PHS（个人手提电话系统）、便携声音处理设备、便携视频处理设备、PDA（个人数字助理）、家用游戏机和便携游戏机之类的信息处理设备。

[0050] 声音拾取部分 110 拾取声音并执行所拾取声音的离散量化。声音拾取部分 110 包括物理上相互分开的两个或多个拾取部分（例如麦克风）。在图 3 的示例中，声音拾取部分 110 包括两个拾取部分，一个用于拾取左声 L 一个用于拾取右声 R。声音拾取部分 110 输出离散量化的左声 L 和右声 R，作为声音判定部分 120 和声源分离部分 140 的输入声音。

[0051] 声音判定部分 120 判定来自声音拾取部分 110 的输入声音是否包括从声音记录部分 10 附近发出的邻近声 (proximity sound)（第一声音），如操作者的语音或由操作者的动作引起的噪声。下面参考图 4 来描述声音判定部分 120 的详细配置。

[0052] 图 4 是示出声音判定部分 120 的配置的功能框图。声音判定部分 120 包括具有音量检测器 124、平均音量检测器 126 和最大音量检测器 128 的音量检测部分 122，具有频谱检测器 132 和品质检测器 134 的品质检测部分 130，距离 / 方向估计器 136，以及操作者声音估计器 138。在图 4 中，为了使示图清晰，左声 L 和右声 R 一起被示为输入声音。

[0053] 音量检测器 124 检测每个预定的帧周期（例如，几十微秒）内给定的输入声音的音量线 (volume string)（幅度），并将检测到的输入声音的音量线输出到平均音量检测器 126、最大音量检测器 128、品质检测器 134 和距离 / 方向估计器 136。

[0054] 平均音量检测器 126 例如基于从音量检测器 124 输入的每帧的音量线来检测每帧输入声音的音量平均值。然后平均音量检测器 126 将检测到的音量平均值输出到品质检测器 134 和操作者声音估计器 138。

[0055] 最大音量检测器 128 例如基于从音量检测器 124 输入的每帧的音量线来检测每帧输入声音的音量最大值。然后最大音量检测器 128 将检测到的音量最大值输出到品质检测器 134 和操作者声音估计器 138。

[0056] 频谱检测器 132 通过例如对输入声音执行 FFT（快速傅立叶变换）来检测输入声音的频域中的每个频谱。然后频谱检测器 132 将检测到的频谱输出到品质检测器 134 和距离 / 方向估计器 136。

[0057] 品质检测器 134 接收输入声音、音量平均值、音量最大值和频谱，基于这些输入来检测关于输入声音是人声的可能性、是音乐的可能性、平稳性 (stationarity)、冲激性 (impulsiveness) 等，并将结果输出到操作者声音估计器 138。是人声的可能性可以是指示输入声音的部分或整体是否与人声相匹配或多接近于人声的信息。是音乐的可能性可以是指示输入声音的部分或整体是否为音乐或多接近于音乐的信息。

[0058] 平稳性指示声音的统计特性不随时间发生重大变化，如空调声。冲激性指示由于聚能集中 (spot energy concentration) 造成的高度嘈杂的属性，如碰撞声或爆破声。

[0059] 品质检测器 134 例如可基于输入声音的频谱分布和人声的频谱分布之间的匹配度来检测是人声的可能性。另外，品质检测器 134 可将每帧的音量最大值与另一帧的音量最大值相比较，并判定当音量最大值较大时冲激性较高。

[0060] 品质检测器 134 可利用诸如过零点或 LPC（线性预测编码）分析之类的信号处理技术来分析输入声音的品质。由于过零点技术检测输入声音的基本周期，因此品质检测器 134 可基于检测到的基本周期是否在人声的基本周期（例如 100 到 200Hz）内来检测是人声的可能性。

[0061] 距离 / 方向估计器 136 充当位置信息计算部分，该部分接收输入声音、输入声音的音量线、输入声音的频谱等，并估计诸如输入声音的声源或者输入声音中主要包括的声音的声源的方向信息和距离信息之类的位置信息。距离 / 方向估计器 136 联合使用各种基于输入声音的相位、音量、音量线、过去平均音量和最大音量等等的声源位置信息估计方法，从而甚至在声音记录设备 10 主体的回响或声音反射影响严重时也全面估计声源位置。下面参考图 5 至 8 描述距离 / 方向估计器 136 中使用的估计方向信息和距离信息的方法的示例。

[0062] 图 5 是示出基于两个输入声音之间的相位差来估计输入声音的声源位置的方法的图示。若声源被假定为点声源，则可以测量分别到达组成声音拾取部分 110 的麦克风 M1 和麦克风 M2 的输入声音的相位以及输入声音之间的相位差。另外，利用输入声音的相位差、频率 f 和声速 c，可以计算麦克风 M1 到输入声音的声源位置的距离与麦克风 M2 到输入声音的声源位置的距离之差。声源存在于距离差恒定的点的集合中。距离差恒定的这些点的集合由双曲线表示。

[0063] 例如，假设麦克风 M1 位于 (x1,0)，假设麦克风 M2 位于 (x2,0)（利用该假设保持一般性）。若将要计算的声源位置的点的集合中的一点为 (x,y) 且上述距离差为 d，则获得下述表达式 1。

[0064] [表达式 1]：

[0065] $$\sqrt{(x-x_1)^2+y^2}-\sqrt{(x-x_2)^2+y^2}=d$$

[0066] 表达式 1 可被扩展成表达式 2，表达式 2 可被整理成表示双曲线的表达式 3。

[0067] [表达式 2]：

[0068] $$\{(x-x_1)^2+2y^2+(x-x_2)^2-d^2\}^2 = 4\{(x-x_1)^2+y^2\}\{(x-x_2)^2+y^2\}$$

[0069] [表达式 3]：

[0070] $$\frac{\left(x-\frac{x_1+x_2}{2}\right)^2}{\left(\frac{d}{2}\right)^2}-\frac{y^2}{\left(\frac{1}{2}\right)^2}=1$$

[0071] 距离 / 方向估计器 136 可以基于分别由麦克风 M1 和麦克风 M2 拾取的输入声音之间的音量差来判定声源位于麦克风 M1 和麦克风 M2 中的哪个附近。例如，距离 / 方向估计器 136 可以判定声源存在于如图 5 所示的麦克风 M2 附近的双曲线 1 上。

[0072] 用于计算相位差的输入声音的频率 f 需满足表达式 4 表示的条件，该表达式与麦克风 M1 和麦克风 M2 之间的距离有关。

[0073] [表达式 4] ：

[0074] $$f<\frac{c}{2d}$$

[0075] 图 6 是示出基于三个输入声音之间的相位差来估计输入声音的声源位置的方法的图示。假设组成声音拾取部分 110 的麦克风 M3、麦克风 M4 和麦克风 M5 按图 6 所示设置。若到达麦克风 M5 的输入声音的相位与到达麦克风 M3 和麦克风 M4 的输入声音的相位相比滞后，则距离 / 方向估计器 136 判定声源位于麦克风 M5 关于连接麦克风 M3 和麦克风 M4 的直线 1 的相反一侧 (深度判定) 。

[0076] 另外，距离 / 方向估计器 136 可以基于分别到达麦克风 M3 和麦克风 M4 的输入声音之间的相位差来计算上面可能存在声源的双曲线 2，还可基于分别到达麦克风 M4 和麦克风 M5 的输入声音之间的相位差来计算上面可能存在声源的双曲线 3。结果，距离 / 方向估计器 136 可以将双曲线 2 和双曲线 3 的交点 P1 估计为声源位置。

[0077] 图 7 是示出基于两个输入声音的音量来估计输入声音的声源位置的方法的图示。若声源被假定为点声源，则在给定点处测量到的音量根据平方反比定律 (inverse square law) 与距离的平方成反比。假设麦克风 M6 和麦克风 M7 如图 7 所示组成声音拾取部分 110，则分别到达麦克风 M6 和麦克风 M7 的声音的音量比恒定的点的集合由圆表示。距离 / 方向估计器 136 可以根据从音量检测器 124 输入的音量值来计算音量比，然后计算上面存在声源的圆的半径和圆心。

[0078] 如图 7 所示，当麦克风 M6 位于 (x3,0) 且麦克风 M7 位于 (x4,0) 时 (利用该假设保持一般性) ，若将要计算的声源位置的点的集合中的一点为 (x,y) ，则分别从麦克风 M6 和 M7 到声源的距离 r1 和 r2 如下述表达式 5 所示。

[0079] [表达式 5] ：

[0080] $$r_1=\sqrt{(x-x_3)^2+y^2} \qquad r_2=\sqrt{(x-x_4)^2+y^2}$$

[0081] 表达式 6 由平方反比定律获得。

[0082] [表达式 6] ：

[0083] $$\frac{1}{r_1^2}:\frac{1}{r_2^2}=const.$$

[0084] 利用正常数 d(例如 4) ，表达式 6 可转化为表达式 7。

[0085] [表达式 7] ：

[0086] $$\frac{r_2^2}{r_1^2}=d$$

[0087] 若将 r1 和 r2 代入表达式 7 并整理，则获得下述表达式 8。

[0088] [表达式 8]：

[0089] $$\frac{(x-x_4)^2+y^2}{(x-x_3)^2+y^2}=d$$

[0090] $$\left(x-\frac{x_4-dx_3}{1-d}\right)^2+y^2=\frac{d(x_4-x_3)^2}{(1-d)^2}$$

[0091] 根据表达式 8，距离 / 方向估计器 136 估计声源存在于圆心坐标由表达式 9 来表示且半径由表达式 10 来表示的圆 1 上，如图 7 所示。

[0092] [表达式 9]：

[0093] $$\left(\frac{x_4-dx_3}{1-d},0\right)$$

[0094] [表达式 10]：

[0095] $$\left|\frac{x_4-x_3}{1-d}\right|\sqrt{d}$$

[0096] 图 8 是示出基于三个输入声音的音量来估计输入声音的声源位置的方法的图示。假设组成声音拾取部分 110 的麦克风 M3、麦克风 M4 和麦克风 M5 按图 8 所示设置，若到达麦克风 M5 的输入声音的相位与到达麦克风 M3 和麦克风 M4 的输入声音的相位相比滞后，则距离 / 方向估计器 136 判定声源位于麦克风 M5 关于连接麦克风 M3 和麦克风 M4 的直线 2 的相反一侧（深度判定）。

[0097] 另外，距离 / 方向估计器 136 可以基于分别到达麦克风 M3 和麦克风 M4 的输入声音的音量比来计算上面可能存在声源的圆 2，还可基于分别到达麦克风 M4 和麦克风 M5 的输入声音的音量比来计算上面可能存在声源的圆 3。结果，距离 / 方向估计器 136 可以将圆 2 和圆 3 的交点 P2 估计为声源位置。通过使用 4 个或更多麦克风，距离 / 方向估计器 136 可以取得包括声源的空间布局在内的更精确的估计。

[0098] 距离 / 方向估计器 136 如上所述基于输入声音的音量比或相位差来估计输入声音的声源位置，然后将有关估计的声源的方向信息和距离信息输出到操作者声音估计器 138。下面的表 1 总结了上述音量检测部分 122、品质检测部分 130 和距离 / 方向估计器 136 中每个元件的输入和输出。

[0099] [表 1]：

[0100]

模块	输入	输出
音量检测器	输入声音	帧中音量线（幅度）
平均音量检测器	帧中音量线（幅度）	音量平均值
最大音量检测器	帧中音量线（幅度）	音量最大值

[0101]

频谱检测器	输入声音	频谱
品质检测器	输入声音 音量平均值 音量最大值 频谱	人声可能性 音乐可能性 是否平稳 冲激性
距离/方向估计器	输入声音 帧中音量线（幅度） 频谱	方向信息 距离信息

[0102] 若从多个声源发出的声音叠加到输入声音上，则距离 / 方向估计器 136 很难估计输入声音中主要包括的声音的精确声源位置。但是，距离 / 方向估计器 136 可以估计输入声音中主要包括的声音的声源位置附近的位置。另外，由于估计的声源位置可被用作声源分离部分 140 中声音分离的默认值，因此即使距离 / 方向估计器 136 估计的声源位置有误差，声音记录设备 10 也可以执行希望的操作。

[0103] 参考回图 4，下面进一步描述声音判定部分 120 的配置。操作者声音估计器 138 基于输入声音的音量、品质和位置信息中的至少一个来全面判定输入声音是否包括从位于声音记录设备 10 附近的特定声源发出的邻近声，如来自操作者的声音或由操作者的动作引起的噪声。当操作者声音估计器 138 判定输入声音包括邻近声时，它还充当将指示输入声音包括邻近声的信息（操作者声音存在信息）、距离 / 方向估计器 136 估计的位置信息等输出到声源分离部分 140 的声音判定部分。

[0104] 具体而言，当距离 / 方向估计器 136 估计输入声音的声源位于拍摄视频的拍摄部分（未示出）的拍摄方向后方且输入声音匹配或近似于人声时，操作者声音估计器 138 可判定输入声音包括邻近声。如图 9 所示，操作者从拍摄部分的拍摄方向后方（通常情况下是取景器的左后方）操纵声音记录设备 10（习惯右手的人正常拍摄而非自拍期间）。

[0105] 因此，若输入声音的声源位置在拍摄部分的拍摄方向后方且输入声音匹配或近似于人声，则操作者声音估计器 138 可以判定输入声音主要包括作为邻近声的操作者声音。因此可以获得操作者的音量比例被声音混合部分 150 减少的混合声，如后所述。

[0106] 另外，若输入声音的声源位置在离声音拾取位置规定距离之内（例如，声音记录设备 10 附近，离声音记录设备 10 一米以内之类），输入声音包括冲激声，并且输入声音的音量高于过去平均音量，则操作者声音估计器 138 可判定输入声音包括从特定声源发出的邻近声。当声音记录设备 10 的操作者操纵声音记录设备 10 的按钮或改变声音记录设备 10 的手持方式时，很可能出现啪嗒或砰之类的冲激声。由于这类冲激声在声音记录设备 10 中生成，因此很可能以相对较高的音量被拾取。

[0107] 因此，若输入声音的声源位置在离声音拾取位置规定距离之内，输入声音包括冲激声，且输入声音的音量高于过去的平均音量，则操作者声音判定估计器 138 可以判定输入声音主要包括作为邻近声的由操作者的动作引起的噪声。因此可以获得由操作者的动作造成的噪声的音量比例被声音混合部分 150 减少的混合声，如后所述。

[0108] 下面的表 2 总结了被输入到操作者声音估计器 138 的信息以及操作者声音估计器 138 做出的判定结果的示例。可以通过联合使用邻近度传感器、温度传感器等来增强操作者声音估计器 138 中的判定精度。

[0109]

[表 2]：

操作者声音估计器的输入									判定结果
音量		品质					方向和距离		
音量	平均音量	最大音量	人声可能性	音乐可能性	平稳与否	冲激性	方向	距离	
高	高于过去平均音量	高	高	低	非平稳	中	后向	近	操作者声音
高	高于过去平均音量	高	低	低	非平稳	高	全向	近	操作噪声
低	低于过去平均音量	低	中	中	平稳	低	不明	远	平稳的环境声
低	相对低于过去平均音量	中	低	低	非平稳	高	全向	远	冲击性环境声
中	相对高于过去平均音量	中到高	中	中	非平稳	中	前向	近到远	对象声音

[0110] 参考回图 3，下面进一步描述声音记录设备 10 的配置。当声源分离部分 140 从声音判定部分 120 接收操作者声音存在信息时，它基于从声音判定部分 120 输入的声源位置

信息来将从声音拾取部分 110 输入的输入声音分离成诸如操作者声音之类的邻近声和不同于邻近声的诸如对象声音之类的拾取目标声（第二声音）。因此，声源分离部分 140 输出双倍数目的输入声音。图 3 示出声源分离部分 140 接收左声 L 和右声 R 作为输入声音，输出左邻近声 L1 和右邻近声 R1 作为邻近声，还输出左拾取目标声 L2 和右拾取目标声 R2 作为拾取目标声的示例。

[0111] 具体而言，声源分离部分 140 充当利用使用 ICA（独立分量分析）的技术、使用声音的时间 - 频率分量之间的微小重叠的技术等等来根据声源分离声音的声音分离部分。

[0112] 声音混合部分 150 以混合声中邻近声的音量比例低于输入声音中邻近声的音量比例的方式来混合从声源分离部分 140 输入的邻近声和拾取目标声。在该配置中，当输入声音中从特定声源发出的邻近声的音量过高时，声音混合部分 150 可以生成拾取目标声的音量比例高于输入声音中拾取目标声的音量比例的混合声。因此，声音记录设备 10 可以避免拾取目标声被不当地掩埋在邻近声中。

[0113] 声音混合部分 150 混合左邻近声 L1 和左拾取目标声 L2 以生成混合左声 L，并混合右邻近声 R1 和右拾取目标声 R2 以生成混合右声 R。然后声音混合部分 150 将混合左声 L 和混合右声 R 作为混合声输出到记录部分 160。

[0114] 声音混合部分 150 可根据被声源分离部分 140 分离的邻近声和拾取目标声的平均音量比计算适当的混合比，并以算出的混合比来混合邻近声和拾取目标声。另外，声音混合部分 150 可在当前帧和前一帧之间的混合比之差不超过预定限度的范围内改变每帧中将使用的混合比。

[0115] 记录部分 160 将从声音混合部分 150 输入的混合声记录到存储部分 170 中。存储部分 170 可以是诸如 EEPROM（电可擦除可编程只读存储器）和 EPROM（可擦除可编程只读存储器）之类的非易失性存储器，诸如硬盘和扁圆状磁盘之类的磁盘，诸如 CD-R（可录致密盘）/RW（可改写）、DVD-R（可录数字万用盘）/RW/+R/+RW/RAM（随机存取存储器）和 BD（蓝光盘（注册商标））-R/BD-RE 之类的光盘，或者诸如 MO（磁光）盘之类的存储器。存储部分 170 还可存储对象的视频数据。

[0116] 如上所述，在本实施例的声音记录设备 10 中，记录部分 160 将邻近声的音量比例低于输入声音中邻近声的音量比例的混合声记录到存储部分 170 中。这允许重放设备重放适当调整了邻近声的音量比例的混合声，而不需在重放设备上安装专门的音量校正功能。

[0117] 前面描述了根据本实施例的声音记录设备 10 的配置。现在参考图 10，在下文中描述在本实施例的声音记录设备 10 中执行的声音处理方法。

[0118] 图 10 是示出根据本实施例在声音记录设备 10 中执行的声音处理方法的流程的流程图。在声音记录设备 10 中，声音拾取部分 110 首先拾取声音 (S210)。若没有输入声音，则过程结束。另一方面，若存在输入声音，则距离 / 方向估计器 136 估计诸如发出全部或部分输入声音的声源的距离或方向之类的位置信息 (S230)。

[0119] 然后，操作者声音估计器 138 判定输入声音是否包括诸如从操作者发出的声音或由操作者的动作引起的噪声之类的邻近声 (S240)。若操作者声音估计器 138 判定输入声音包括邻近声，则声源分离部分 140 将输入声音分离成邻近声和不同于邻近声的拾取目标声 (S250)。

[0120] 之后，声音混合部分 150 以预定的混合比来混合被声源分离部分 140 分离的邻近

声和拾取目标声，从而生成混合声 (S260)。在步骤 S260 之后或者当步骤 S240 判定输入声音不包括诸如从操作者发出的声音或由操作者的动作引起的噪声之类的邻近声时，记录部分 160 将混合声或者输入声音存储到存储部分 170 中。

[0121] 如上所述，在本实施例的声音记录设备 10 中，声源分离部分 140 基于距离 / 方向估计器 136 估计的输入声音的声源位置信息来分离输入声音中包括的从特定声源发出的邻近声，并且声音混合部分 150 以混合声中邻近声的音量比例低于输入声音中邻近声的音量比例的方式来混合邻近声和作为输入声音包括的其他声音的拾取目标声。

[0122] 因此，若输入声音中从特定声源发出的邻近声的音量过高，则声音混合部分 150 可以生成拾取目标声的音量比例高于输入声音中拾取目标声的音量比例的混合声。结果，声音记录设备 10 可以抑制邻近声的相对音量从而避免拾取目标声被不当地掩埋在邻近声中。另外，声音记录设备 10 可以记录输入声音中包括的诸如噪声或从操作者发出的声音之类的邻近声的效果被减轻或消除的高品质混合声。

[0123] 声音记录设备 10 可以记录邻近声的音量比例低于输入声音中邻近声的音量比例的混合声。这允许重放设备重放适当调整了邻近声的音量比例的混合声，而不需在重放设备上安装专门的音量校正功能。

[0124] 另外，由于本实施例的声音记录设备 10 可以通过软件处理输入声音并记录调整了邻近声和拾取目标声的音量比的混合声，因此可以缩小诸如麦克风数目之类的硬件规模。

[0125] 【第二实施例】

[0126] 下面描述根据本发明第二实施例的声音重放设备 11。本实施例的声音重放设备 11 能够重放调整了预存声音中包括的邻近声的音量比例的混合声。下面参考图 11 来描述声音重放设备 11 的配置。

[0127] 图 11 是示出根据本实施例的声音重放设备 11 的配置的功能框图。本实施例的声音重放设备 11 包括声音判定部分 120、声源分离部分 140、声音混合部分 150、存储部分 172、再现部分 174 和声音输出部分 180。在对本实施例的描述中，将不重复描述与第一实施例中描述的配置实际上相同的配置，这里主要描述与第一实施例的配置不同的配置。

[0128] 存储部分 172 存储记录在具有声音记录功能的给定设备中的声音。再现部分 174 读出存储部分 172 中存储的声音并根据需要对其进行解码。然后，再现部分 174 将存储部分 172 中存储的声音输出到声音判定部分 120 和声源分离部分 140。声音判定部分 120 和声源分离部分 140 将来自再现部分 174 的输出视为输入声音，并执行与第一实施例中描述的处理实际上相同的处理。

[0129] 声音输出部分 180 输出由声音混合部分 150 混合的混合声。声音输出部分 180 例如可以扬声器或耳机。与第一实施例中的存储部分 170 类似，本实施例的存储部分 172 也可以是诸如 EEPROM 和 EPROM 之类的非易失性存储器，诸如硬盘和圆盘状磁盘之类的磁盘，诸如 CD-R/RW、DVD-R/RW/+R/+RW/RAM 和 BD（蓝光盘（注册商标））-R/BD-RE 之类的光盘，或者诸如 MO（磁光）盘之类的存储器。

[0130] 如上所述，在本实施例的声音重放设备 11 中，声音判定部分 120、声源分离部分 140 和声音混合部分 150 基于从再现部分 174 输入的输入声音来生成混合声并将混合声作为再现声音输出。这实现了邻近声的音量比例被适当调整的混合声的重放，而不需在将输

入声音记录到存储部分 172 中的记录设备上安装专门的音量校正功能。这还实现了诸如噪声或从操作者发出的声音之类的邻近声的效果被减轻或消除的高品质混合声的输出。

[0131] 【第三实施例】

[0132] 下面描述根据本发明第三实施例的声音重放设备 12。若对输入声音进行了 AGC(自动增益控制),则本实施例的声音重放设备 12 能够反向校正输入声音中包括的拾取目标声的音量并增强拾取目标声。下面参考图 12 和 13 来描述本实施例的声音重放设备 12 的配置和操作。

[0133] 图 12 是示出根据本实施例的声音重放设备 12 的配置的功能框图。声音重放设备 12 包括声音判定部分 120、声源分离部分 140、声音混合部分 150、存储部分 172、再现部分 174、声音输出部分 180 和音量校正部分 190。在对本实施例的描述中,将不详细描述与第二实施例中描述的配置实际上相同的配置,这里主要描述与第二实施例的配置不同的配置。

[0134] 本实施例的存储部分 172 存储被部分或全部进行了 AGC(音量校正)的声音。AGC 是一种压缩器机制,它的一个目标是通过自动降低过高音量输入的音量级来避免限幅噪声 (clipping noise)。下面参考图 13 来描述被实施了 AGC 的声音的音量。

[0135] 图 13 是以对比的方式示出 AGC 实施之前声音(原始声音)的音量和 AGC 实施之后声音的音量的图示。若 AGC 实施之前声音的音量高于阈值 th,则在设为发动时间 (attack time) 的时间内 AGC 将音量压缩至预定比例。图 13 示出 AGC 实施之前声音的音量在设为发动时间的时间内被压缩至 1/2 至 2/3 的情况。之后,若 AGC 实施以后声音的音量低于阈值 th,则在设为解除时间的时间内解除 AGC。

[0136] 常常是在过高音量级的邻近声从声音记录设备附近输入时,音量超过阈值 th 而导致 AGC 的实施。因此,由来自远处声源的拾取目标声造成 AGC 实施的情况不常有。但是,由于输入声音全部被 AGC 压缩,因此不仅输入声音中的邻近声而且原本低音量级的拾取目标声都被 AGC 实施所压缩。

[0137] 鉴于上述问题,发明了根据本实施例的声音重放设备 12。本实施例的声音重放设备 12 能够利用音量校正部分 190 的功能来在输入声音被实施 AGC 时增强拾取目标声。

[0138] 音量校正部分 190 基于声源分离部分 140 分离的邻近声的音量变化来检测其间施加了 AGC 的发动时间,并扫描在声源分离部分 140 分离的拾取目标声中与发动时间相对应的时间段。虽然拾取目标声可能包括背景环境声、从对象发出的声音等,但是若它只包括背景环境声,则它的音量级近似恒定。因此,音量校正部分 190 可以在拾取目标声的音量以预定级别或更高级别变化的时间段内判定 AGC 被实施。

[0139] 这种情况下,音量校正部分 190 执行反向校正,将该时间段中拾取目标声的音量调整为与先前和后续时间段的音量基本上相等,从而增强拾取目标声。

[0140] 若发动时间和解除时间的估计值以及音量校正部分 190 执行的反向校正度被存储,则当拾取目标声包括从对象发出的声音时,这些值可被有效使用。具体而言,当拾取目标声包括从对象发出的声音时,音量校正部分 190 从邻近声中检测发动时间,并扫描拾取目标声中与发动时间相对应的时间段的全部先前和后续时间段中的音量值。作为扫描结果,若音量值以对应于发动时间或解除时间的时间宽度而变化,则音量校正部分 190 可判定实施了 AGC 并执行反向校正。

[0141] 声音混合部分 150 以邻近声的音量比例被抑制的音量比来混合音量被音量校正

部分 190 反向校正过的拾取目标声和被声源分离部分 140 分离的邻近声，从而生成混合声。

[0142] 如前所述，若由于邻近声音量过高而造成输入声音的音量被整体抑制并且拾取目标声的音量被相应地抑制，则根据本发明第三实施例的声音重放设备 12 可以根据输入声音音量的抑制度来增大拾取目标声的音量，从而避免拾取目标声太小。

[0143] 虽然音量校正部分 190 包括在本实施例的声音重放设备 12 中，但是若音量校正部分 190 包括在第一实施例的声音记录设备 10 中，则可以在对输入声音实施 AGC 时将包括根据 AGC 程度被增强的拾取目标声的混合声记录到存储部分 170 中。

[0144] 本领域技术人员应该理解，取决于涉及要求和其他因素可以想到各种修改、组合、子组合及变更，只要它们落入所附权利要求或其等同物的范围之内。

[0145] 例如，不必以根据流程图所示次序的时间顺序来执行声音记录设备 10 的过程的每一步，并行或单独执行的处理（例如，并行处理或对象处理）可被包括在内。

[0146] 另外，虽然图 3 示出声音判定部分 120 判定声音拾取部分 110 拾取的输入声音是否包括邻近声的情况，但是本发明不限于这种情况。例如，声音判定部分 120 可接收声源分离部分 140 分离的声音，估计分离的声音的声源位置，判定分离的声音是否包括邻近声，并将分离的声音输出到声音混合部分 150。这种情况下，声源分离部分 140 对每个的声源声音进行不带默认值的盲分离。

[0147] 另外，可以创建计算机程序，所述程序使得声音记录设备 10、声音重放设备 11 或声音重放设备 12 中内置的诸如 CPU、ROM 或 RAM 之类的硬件执行与声音记录设备 10、声音重放设备 11 或声音重放设备 12 中每一个的配置相同的功能。另外，可以提供存储这种计算机程序的存储器。还可以通过利用硬件实现声音记录设备 10、声音重放设备 11 或声音重放设备 12 的每个功能框图所示的每个功能块来获得硬件上的处理序列。

[0148] 本发明包含与 2007 年 2 月 15 日向日本专利局递交的日本专利申请 JP2007-035410 相关的主题，其全部内容通过引用方式结合于此。

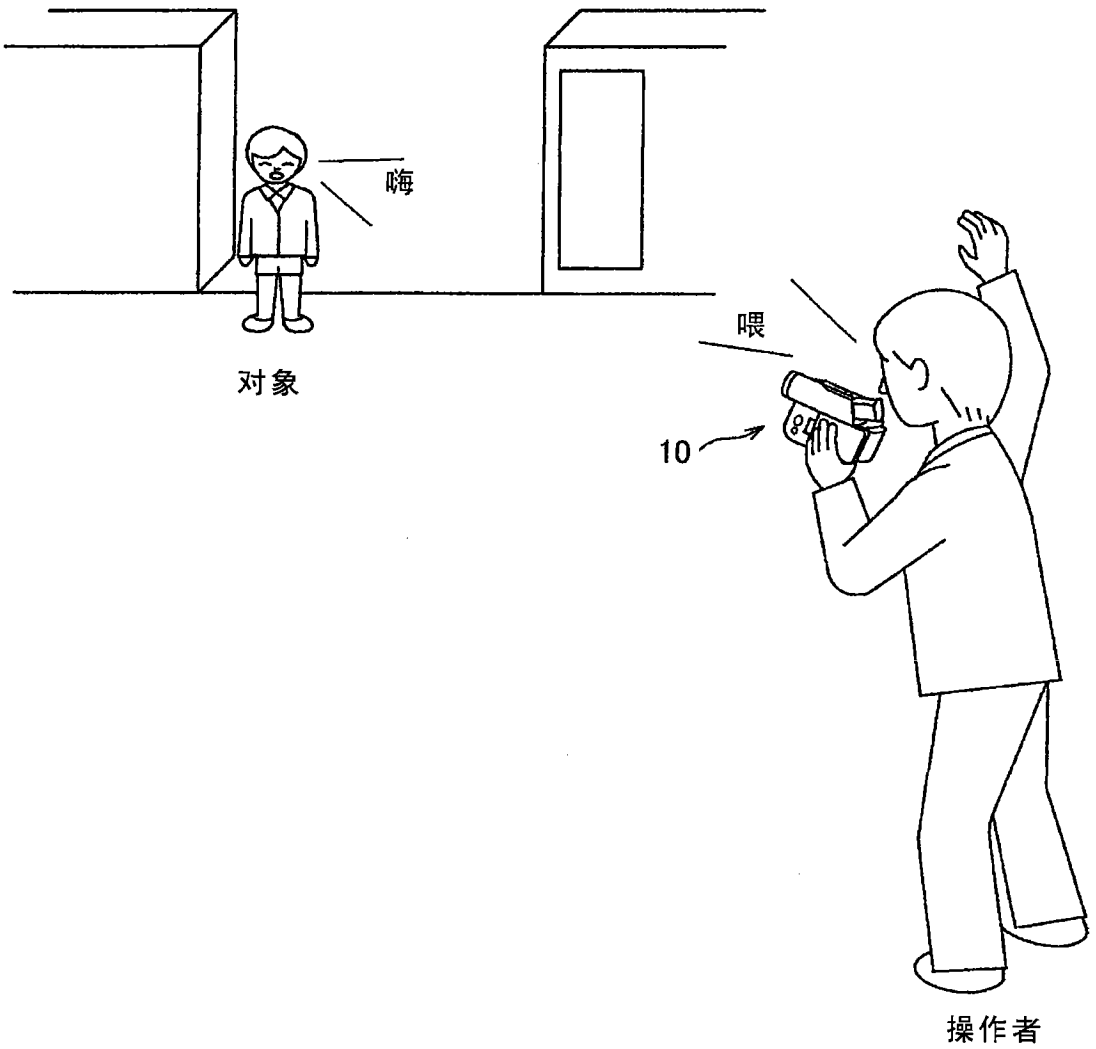


图1

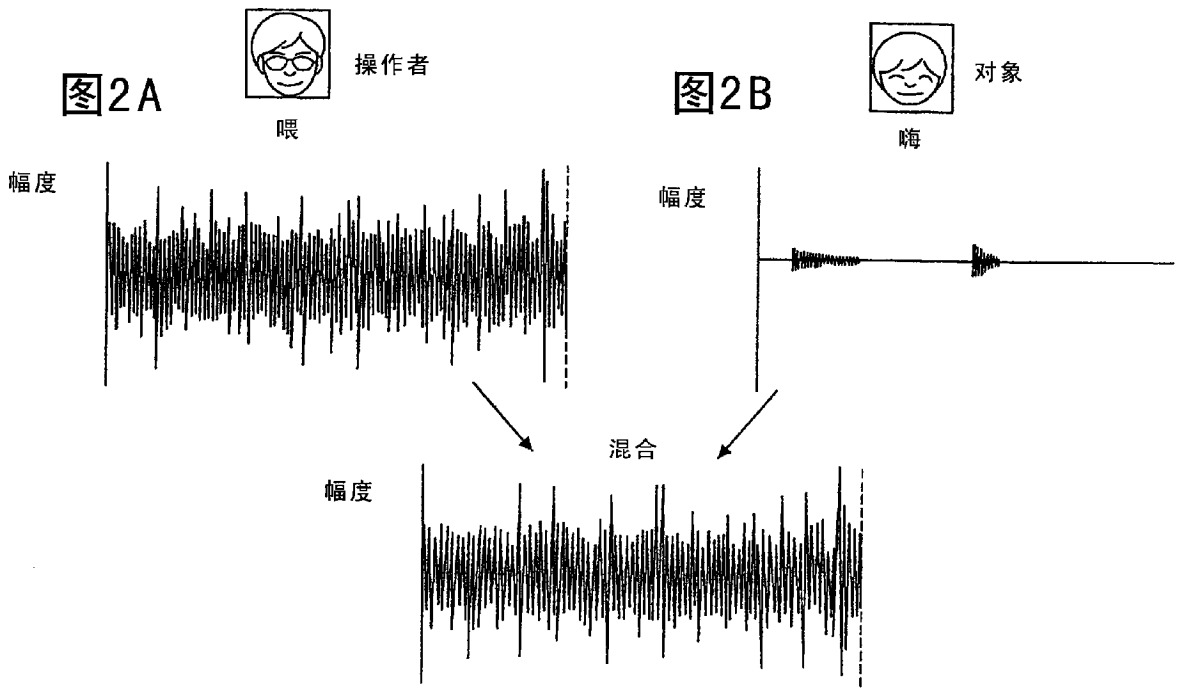


图2A 图2B

图2C

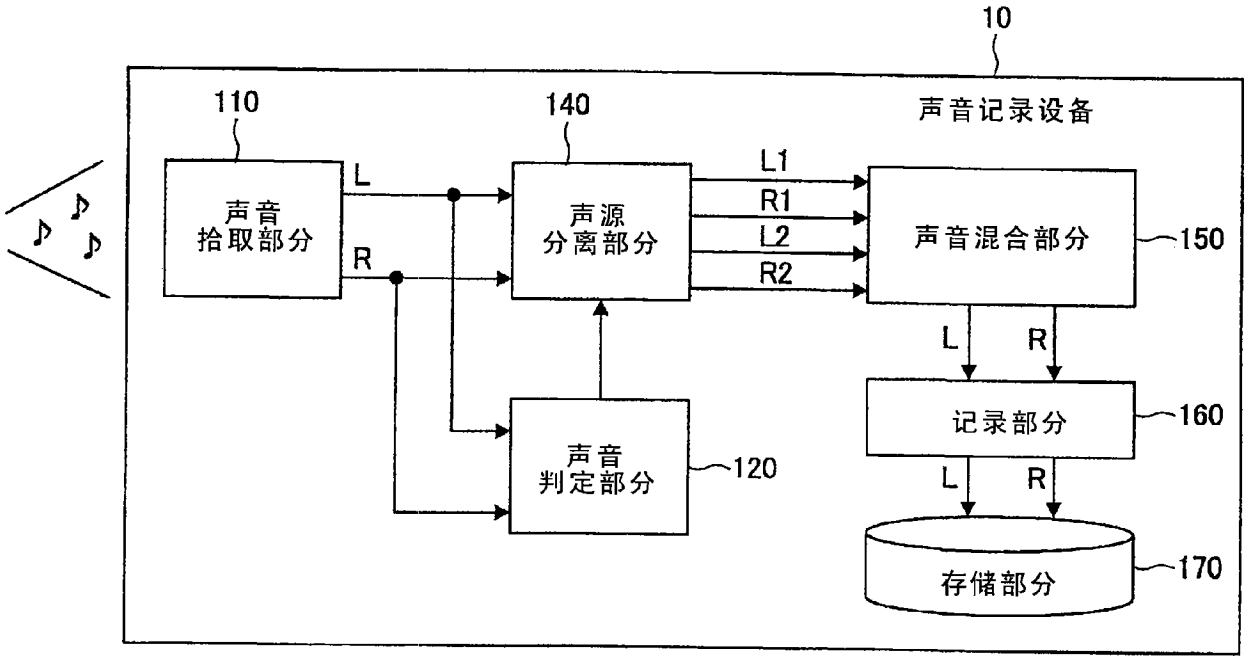

10
声音记录设备
110
声音
拾取部分
L
R
140
声源
分离部分
L1
R1
L2
R2
声音混合部分
150
L
R
记录部分
160
L
R
存储部分
170
声音
判定部分
120

图3

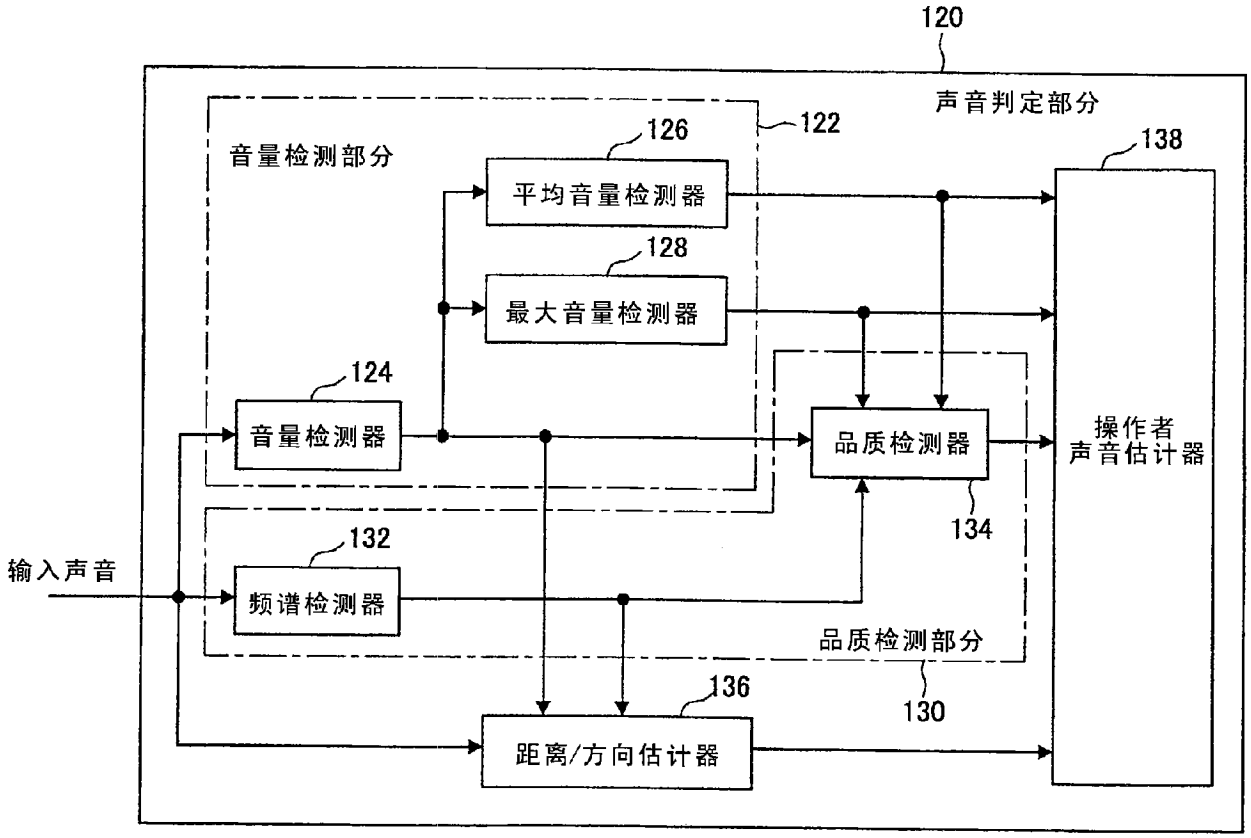

120
声音判定部分
音量检测部分
122
126
平均音量检测器
128
最大音量检测器
124
音量检测器
品质检测器
134
138
操作者
声音估计器
输入声音
132
频谱检测器
品质检测部分
130
136
距离/方向估计器

图4

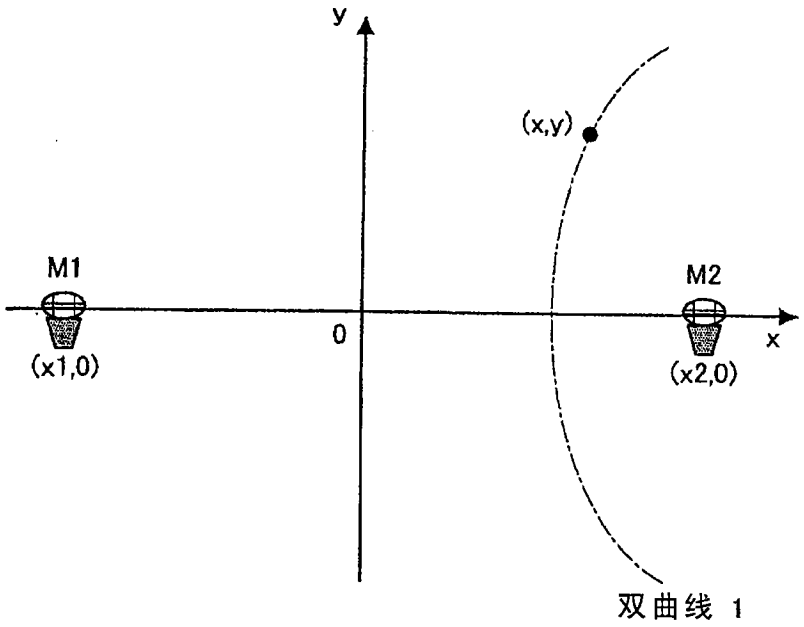


图5

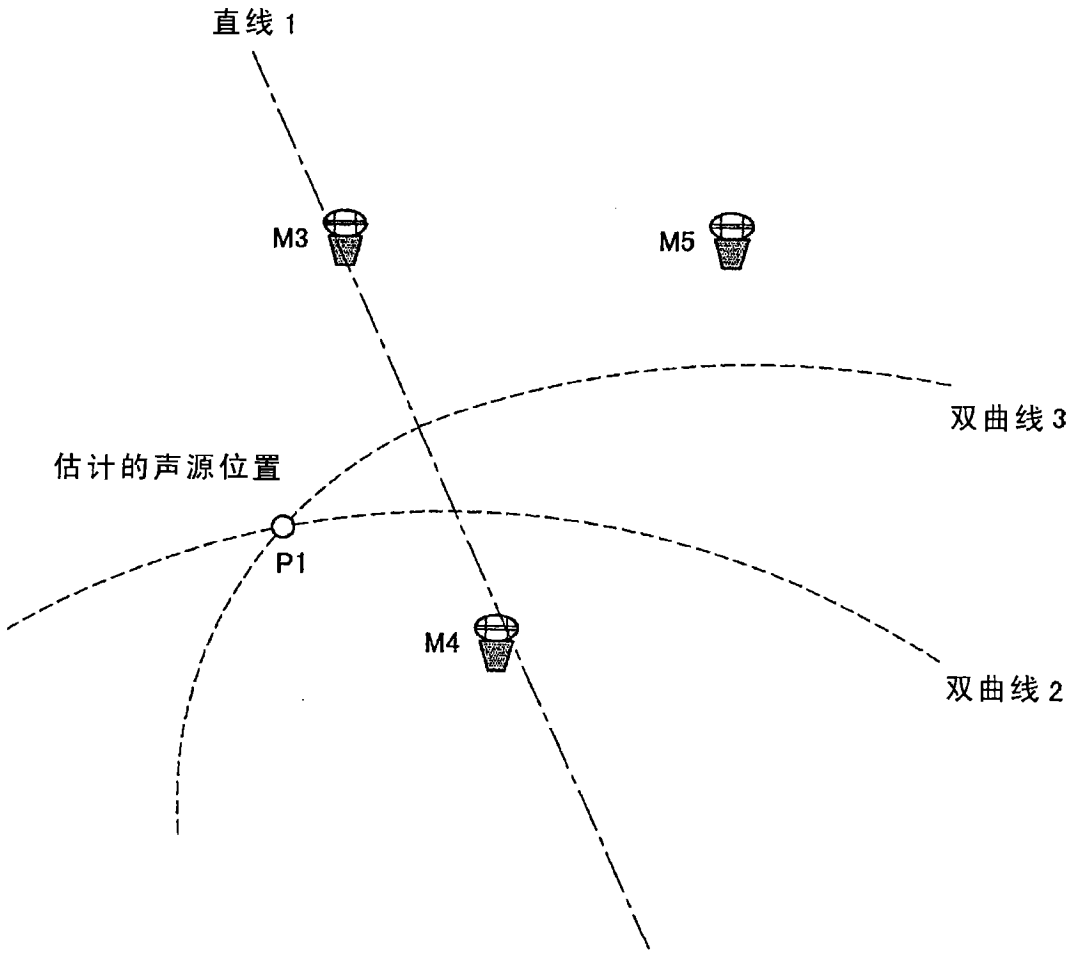


图6

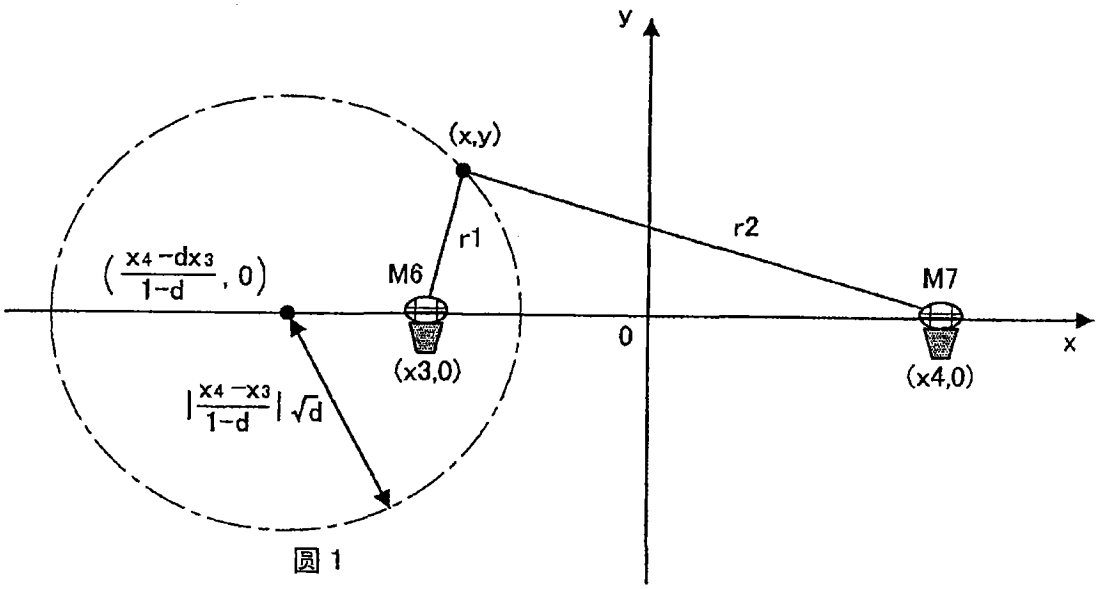

y
(x,y)
r1
r2
$\left(\frac{x_4-dx_3}{1-d},0\right)$
M6
M7
0
x
(x3,0)
(x4,0)
$\left|\frac{x_4-x_3}{1-d}\right|\sqrt{d}$
圆 1

图7

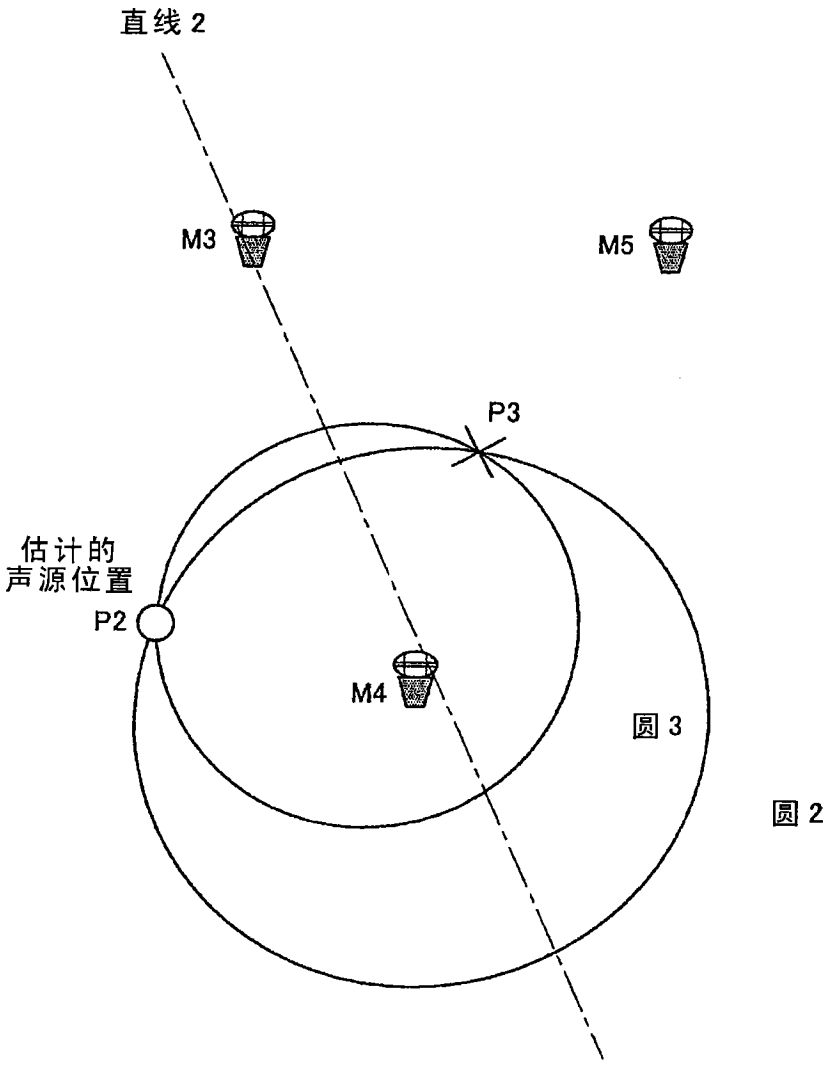


图8

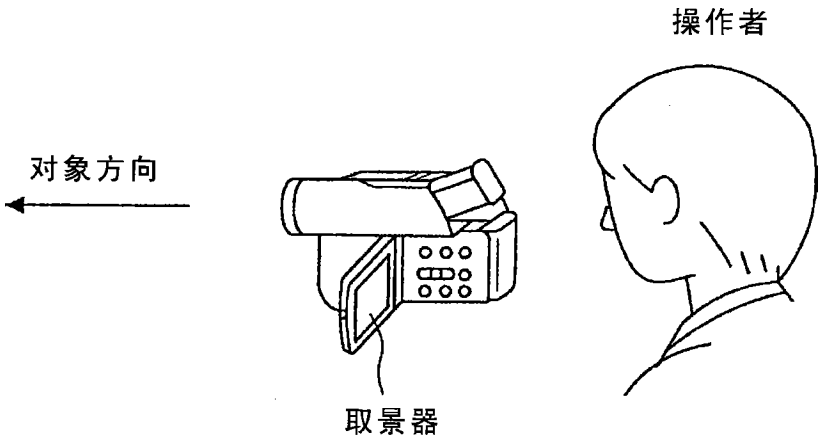


图9

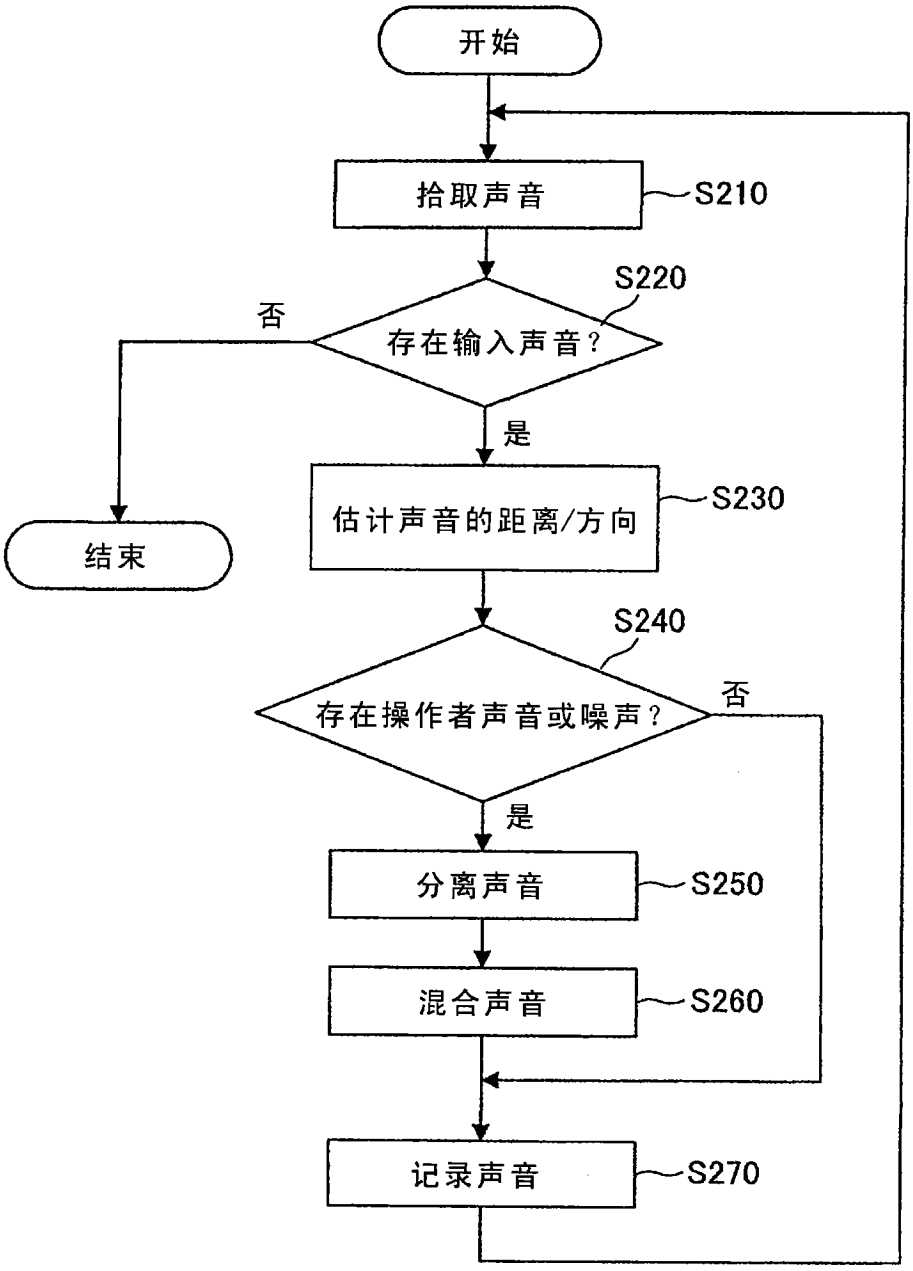

开始
拾取声音
S210
S220
存在输入声音？
否
是
结束
估计声音的距离/方向
S230
S240
存在操作者声音或噪声？
否
是
分离声音
S250
混合声音
S260
记录声音
S270

图10

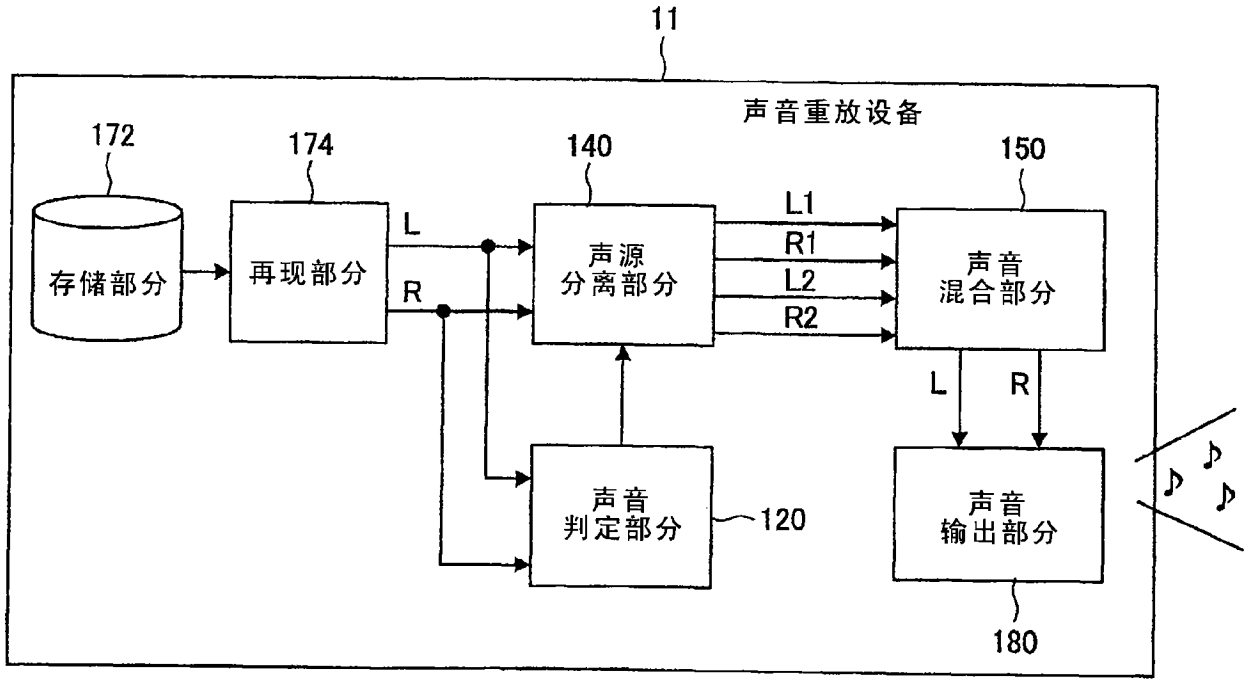

11
声音重放设备
172
存储部分
174
再现部分
L
R
140
声源
分离部分
L1
R1
L2
R2
150
声音
混合部分
L
R
声音
判定部分
120
声音
输出部分
180

图11

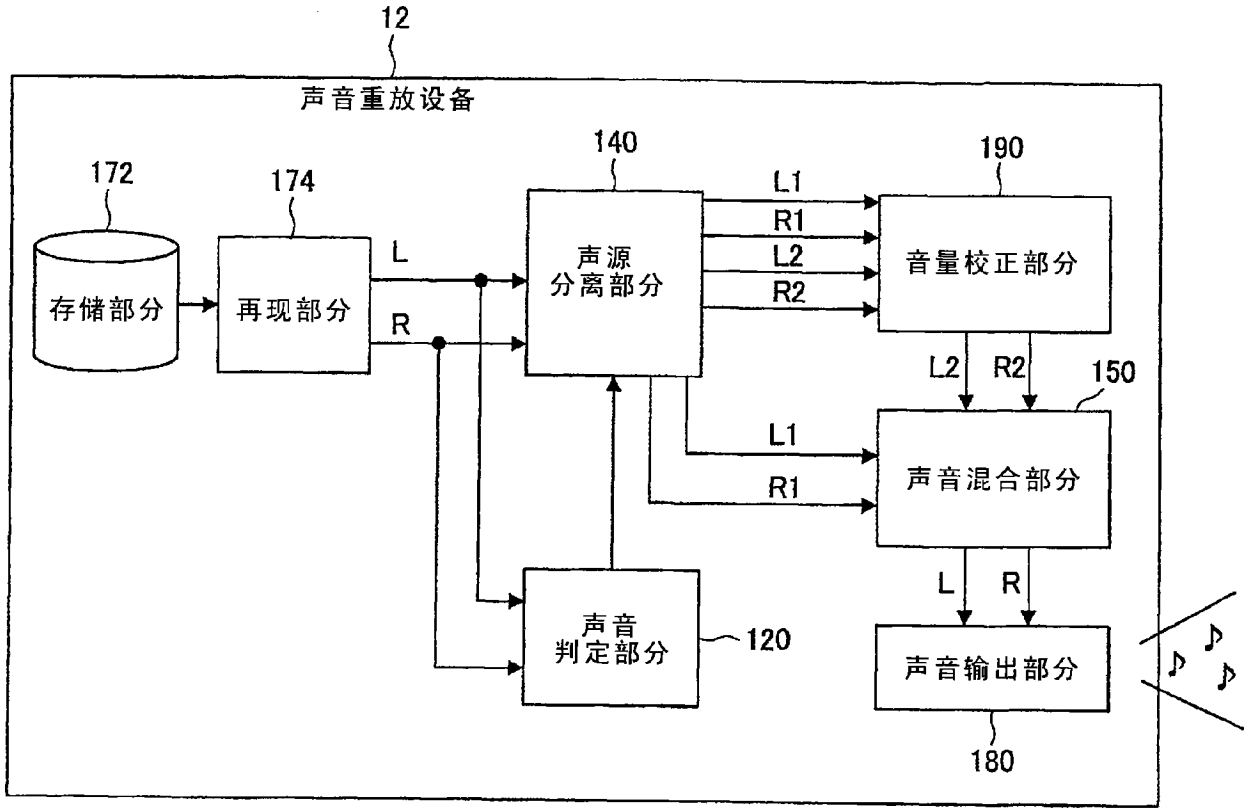

12
声音重放设备
172
存储部分
174
再现部分
L
R
140
声源
分离部分
L1
R1
L2
R2
190
音量校正部分
L2
R2
150
声音混合部分
L1
R1
L
R
声音
判定部分
120
声音输出部分
180

图12

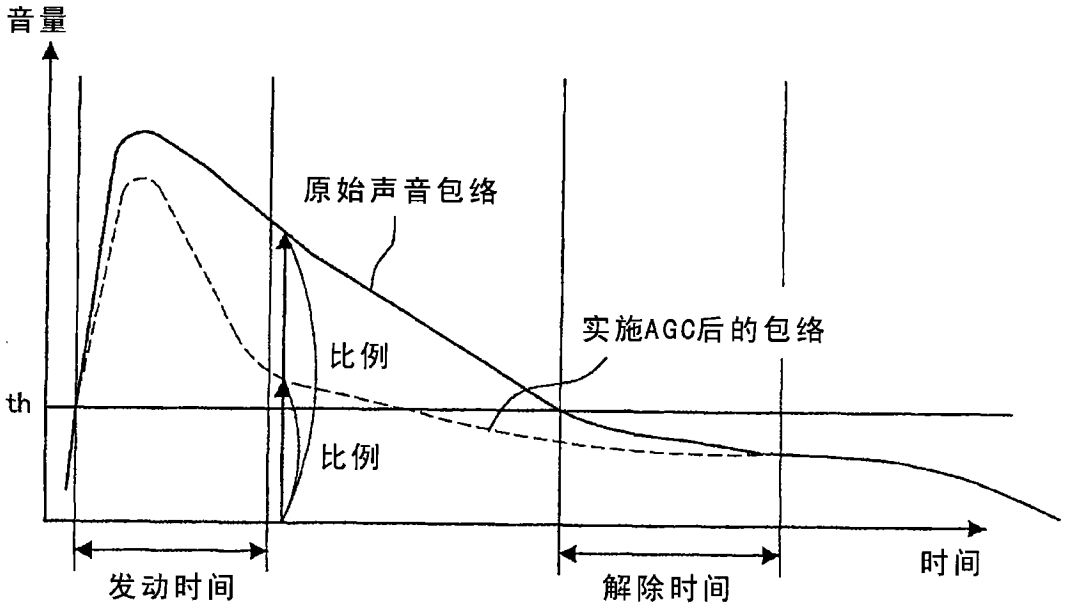


图13